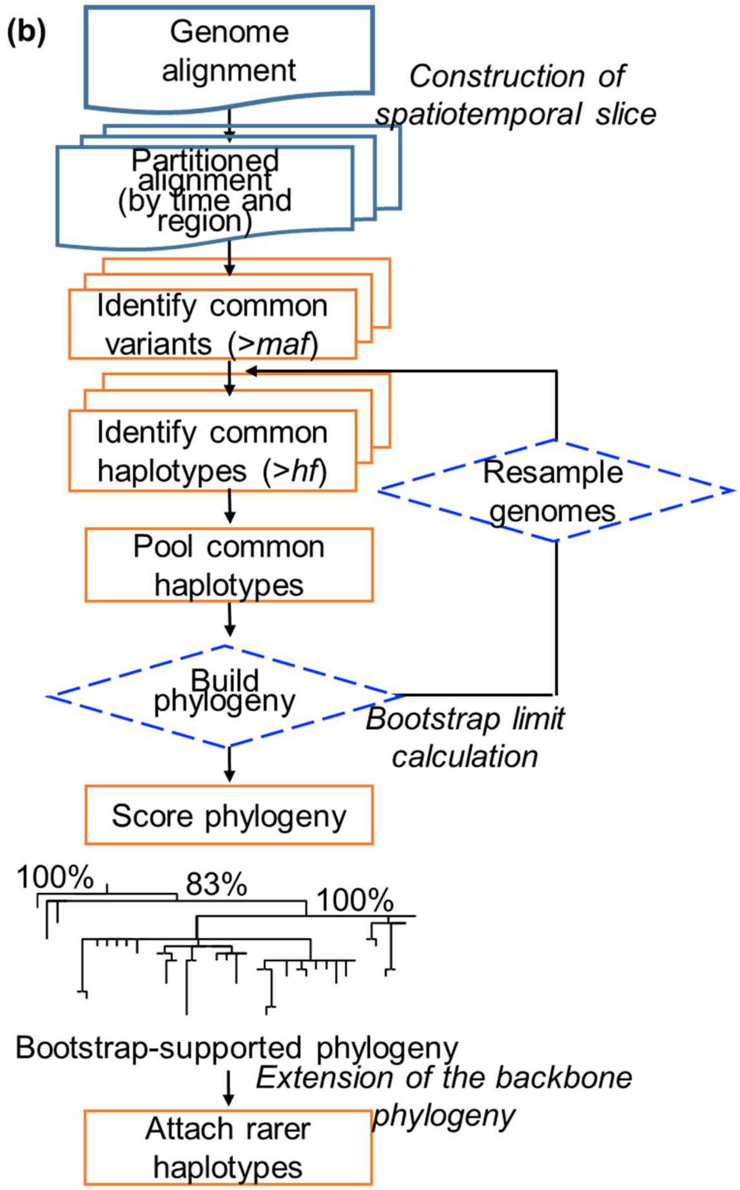

Figure S1. A flowchart of *TopHap* analysis using spatiotemporal information. Input to *TopHap* is an alignment of genome sequences. When the selection of variants and haplotypes is sensitive to spatiotemporal (or other types) annotations of genomes, then variants and haplotypes are identified for each slice separately and then merged before phylogenetic analysis. To compute bootstrap confidence limits, we suggest resampling haplotypes with replacement to build bootstrap replicates for every spatiotemporal slice individually.

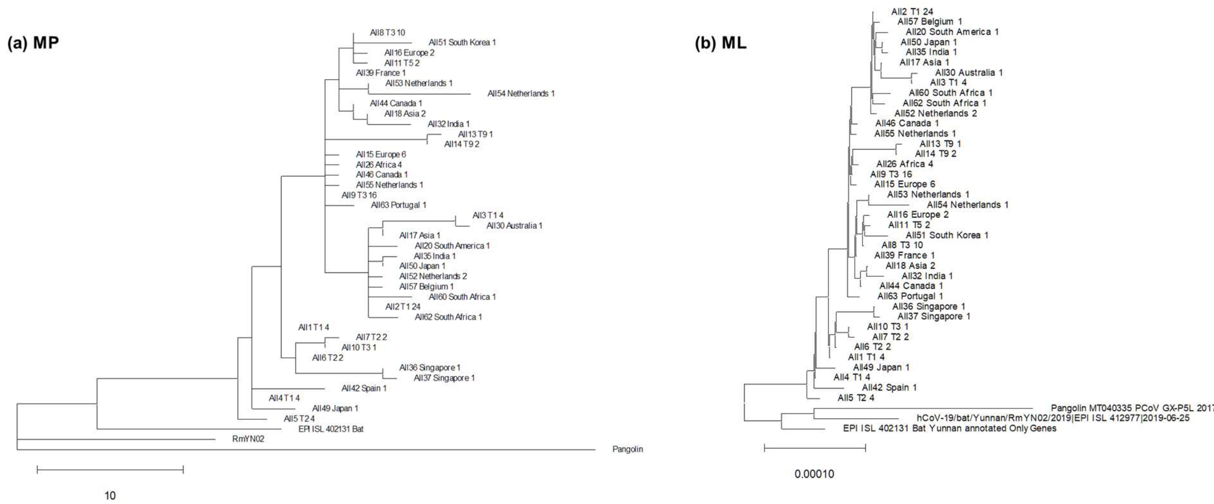

Figure S2. Maximum parsimony (MP; a) and maximum likelihood (ML; b) phylogenies for the 68KG dataset. 5% *maf* and *hf* cutoffs were used to generate an alignment of common haplotypes. MEGA was used to build the MP phylogeny (heuristic search, default options), and RAxML was used to construct the ML phylogeny. In RAxML, the GTR substitution model with GAMMA among-site rate heterogeneity (the number of rate categories equal to 4) was selected. Since the alignment contains only variable sites, ascertainment bias correction was used.

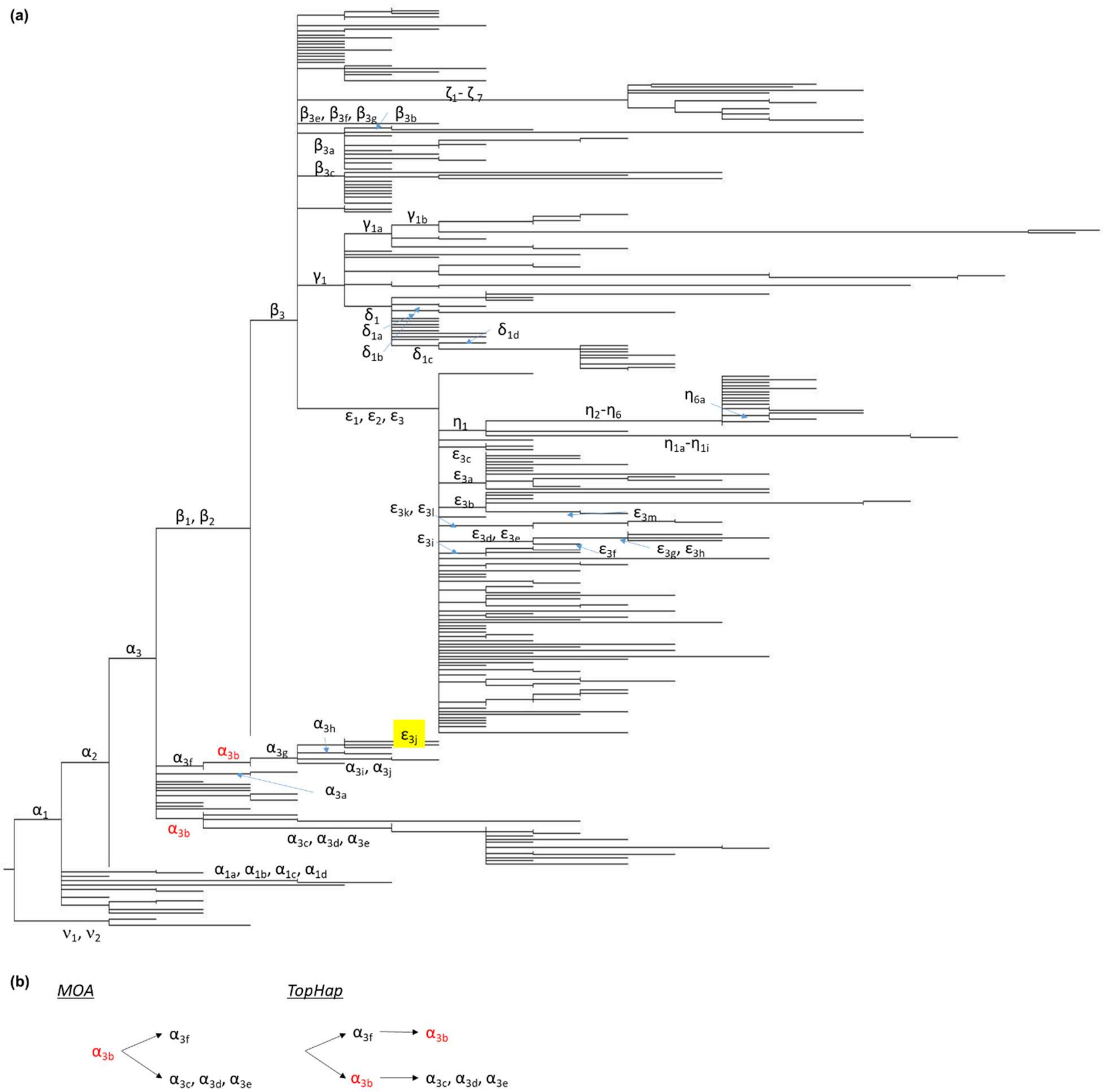

Figure S3. *TopHap* phylogeny using 1% *maf* and *hf thresholds*. The 68KG data was used. (**a**) Mutations ordered using MOA in Kumar et al. (2021) are shown using Greek symbols. The inferred mutation orders by *TopHap* agreed with MOA with the following exceptions. First, the ε_{3j} mutation (yellow box) was predicted to have occurred after the ε_3 mutation by MOA. In the TopHap inference, the haplotype with ε_{3j} mutation does not have β_1 - β_3 and ε_1 - ε_3 mutations, but it has α_{3b}, α_{3f}, α_{3g} mutations, supporting the *TopHap*'s placement. MOA attached the ε_{3j} mutation after ε_3 mutation, probably because ε_{3j} mutation recurrently happened on haplotypes with ε_3 mutation creating rare haplotypes with ε_{3j} and ε_3 mutations. Second, the frequencies of haplotypes with β_{3d} and δ_{1e} mutations were not > 5% in any time slice, so TopHap did not order these mutations. Last, *TopHap* predicted that α_{3b} mutation recurrently occurred while MOA did not (**b**). Note that MOA cannot infer recurrent mutations.

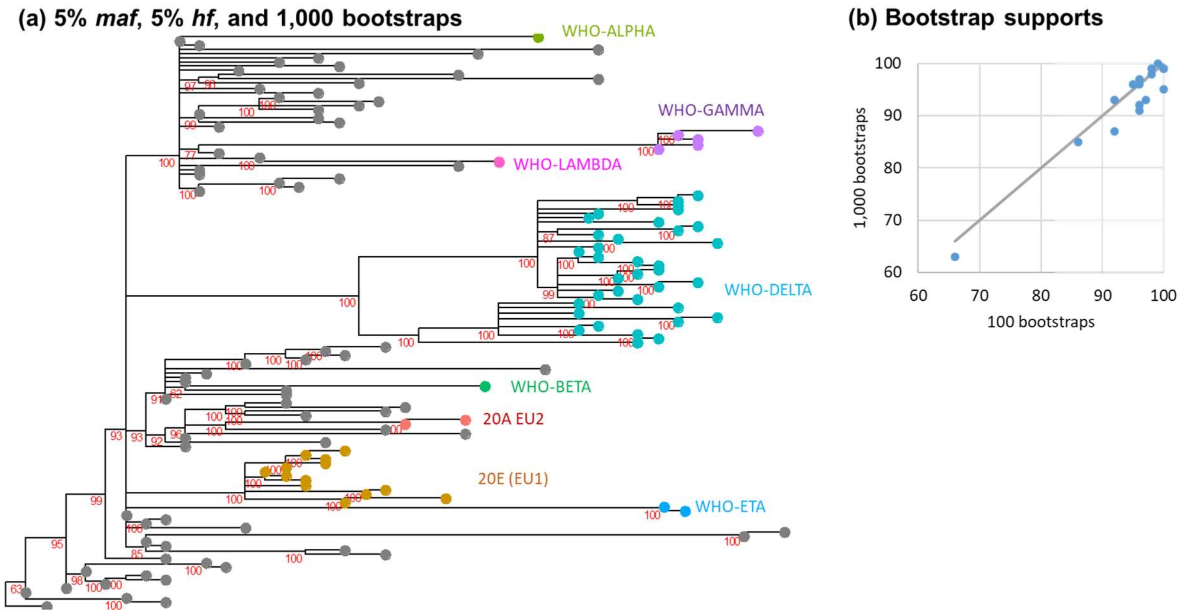

Figure S4. The *TopHap* phylogeny using 1,000 bootstrap replicates. The 1MG data was used with a 5% cutoff for *maf* and *hf*. (**a**) *TopHap* phylogeny. Red numbers near nodes are bootstrap confidence limits derived from bootstrap resampling of genomes. The haplotypes with concerning mutations are indicated using WHO IDs, and 20A EU2 and 20E (EU1) are Nextstrain clade IDs. (**b**) a scatterplot of bootstrap confidence limits obtained using 1,000 and 100 bootstrap replicates. Not shown are estimates in which both analyses produced 100% bootstrap support. Because *TopHap* implementation retains only haplotypes found in all bootstrap replicates, the number of haplotypes decreased to 130 in the bootstrap consensus tree from 1000 replicates.

1% *maf*, 1% *hf*, and 100 bootstraps

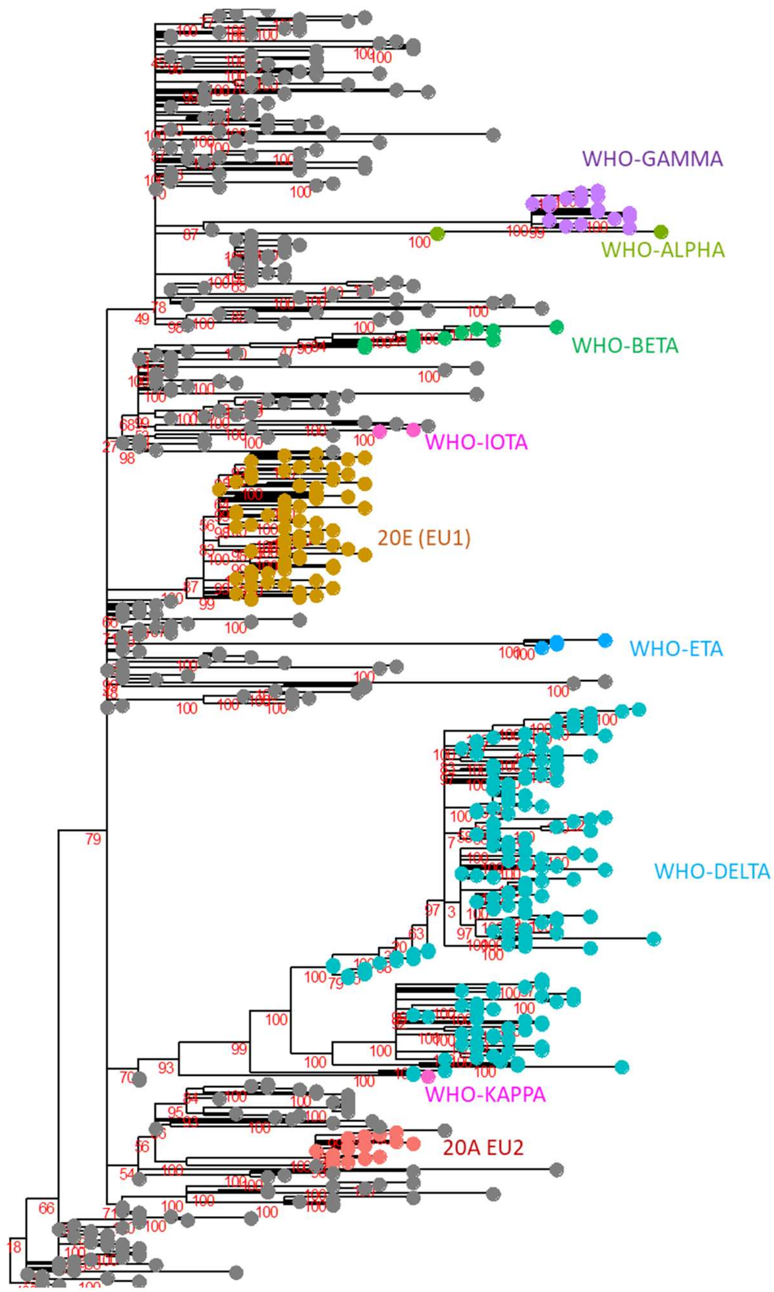

Figure S5. The 1MG *TopHap* phylogeny using 1% *maf* and *hf thresholds*. *TopHap* phylogeny with bootstrap confidence limits (red) is shown. The haplotypes with WHO concerning mutations are indicated by using WHO IDs; 20A EU2 and 20E (EU1) are Nextstrain clade IDs.